

Modellbasierte Ansätze für die computerunterstützte Textinterpretation

Muhr, Thomas; Seibt, Volker

Veröffentlichungsversion / Published Version
Sammelwerksbeitrag / collection article

Empfohlene Zitierung / Suggested Citation:

Muhr, T., & Seibt, V. (1994). Modellbasierte Ansätze für die computerunterstützte Textinterpretation. In A. Boehm, A. Mengel, & T. Muhr (Hrsg.), *Texte verstehen : Konzepte, Methoden, Werkzeuge* (S. 289-314). Konstanz: UVK Univ.-Verl. Konstanz. <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-14646>

Nutzungsbedingungen:

Dieser Text wird unter einer CC BY-NC-ND Lizenz (Namensnennung-Nicht-kommerziell-Keine Bearbeitung) zur Verfügung gestellt. Nähere Auskünfte zu den CC-Lizenzen finden Sie hier:
<https://creativecommons.org/licenses/by-nc-nd/4.0/deed.de>

Terms of use:

This document is made available under a CC BY-NC-ND Licence (Attribution-Non Commercial-NoDerivatives). For more information see:
<https://creativecommons.org/licenses/by-nc-nd/4.0>

Modellbasierte Ansätze für die computerunterstützte Textinterpretation

Thomas Muhr
Technische Universität Berlin

Volker Seibt
S+F Datentechnik Leer

1 Einleitung

In diesem Beitrag wird der Einsatz von Netzwerken für die Textinterpretation und Theoriebildung vorgestellt. Dabei wird auch auf die Grenzen einfacher konzeptueller Netzwerke hingewiesen und ein modellbasierter Ansatz vorgestellt. Die vorgestellten Methoden lassen sich mit dem System ATLAS/ti realisieren und sollen Anregungen geben, die Netzwerkfunktionalität für den Analyseprozeß und die Darstellung von Ergebnissen sinnvoll zu nutzen.

Textinterpretation ist im wesentlichen ein auf neue Erkenntnisse ausgerichtetes "In-Bezug-Setzen" von Texten mit anderen Texten, die Anreicherung vorhandener Texte mit Anmerkungen, sowie die Einordnung von Text(-teilen) in theoretisch relevante Kategorien. Für die Theoriebildung spielt fernerhin die Beziehung *zwischen* diesen Kategorien eine Rolle. Obwohl das Resultat von Textinterpretation in der Regel wieder ein linearer, und damit leicht publizierbarer Text ist, finden sich zumindestens während der Bearbeitung Phasen komplexer Strukturbildungen.

Hier geht es um Möglichkeiten, diese sich im Verlauf der Textinterpretation bildenden Strukturen angemessen zu repräsentieren und den Prozeß der Textinterpretation selbst zu unterstützen. Es ist nicht das Ziel dieses Beitrags, eine aus linguistischer und logischer Sicht vollständige bzw. korrekte Repräsentation für Textwissen vorzustellen oder die bei der Textinterpretation im Kopf des Interpreten stattfindenden Denkprozesse zu beschreiben. Die pragmatisch-"heuristischen" Qualitäten von Netzwerkdarstellungen stehen im Vordergrund, also ihre Möglichkeiten, das Denken des Interpreten anzuregen, den Prozeß der Interpretation auch für andere nachvollziehbar zu gestalten und die Vermittlung der Resultate zu verbessern.

2 Text = Textur?

Obwohl schon im Begriff "Text" (lat. *textus* "Gewebe, Geflecht, Verbindung, Zusammenhang", verwandte Begriffe: Textur, Textilie) Struktur enthalten ist, erscheint er zumindest in seiner physischen Manifestation als bloße sequentielle Abfolge von Zeichen eines Alphabets. Fehlt dem linearen Text etwa der "vertikale Schuß"? Flusser (1992) schreibt treffend:

"Etymologisch bedeutet das Wort 'Text' ein Gewebe und das Wort 'Linie' einen Leinenfaden. Texte aber sind unfertige Gewebe: Sie bestehen aus Linien (der 'Kette') und werden nicht, wie fertige Gewebe, von vertikalen Fäden (dem 'Schuß') zusammengehalten. Die Literatur (das Universum der Texte) ist ein Halbfabrikat."

Jedoch gibt es kaum Texte, die auch in ihrer logischen Repräsentation lediglich sequentiell sind. Querverweise im Text ("siehe S. xy"), ana- oder kataphorische Bezugnahmen auf "entfernte" Textpassagen, jedwede linguistische Maßnahme zur Erhaltung der Kohäsion, Zitate, Fußnoten, Kapiteleinteilungen, im Grunde jeder etwas komplexere Satz, Hervorhebungen durch typographische Merkmale, Layout: all dies sind (intra- und intertextuelle) Konstruktionen, die einen Text zur "Textur" machen und den Leser eines so mit "Schuß" versehenen Textes teilweise, und meist durch den Autor kontrolliert, "aus der Bahn" eines rein sequentiellen Lesens werfen. Neben diesen "Mikro"-Strukturen lenken größere Zusammenhänge, etwa die narrative oder durch Normen oder rhetorische Erfordernisse bedingte Makro-Struktur des Textes den Rezeptionsprozeß. Andere Bezüge sind die eines Textes zur Wirklichkeit, die zum Wissenshintergrund des Lesers und den Intentionen des Autors (falls er welche hatte).

Eco (1990) gibt ein Beispiel für einen ausschließlich linear manifestierten Text ohne Inhaltsbezug (aus *Der große Lalula* von Christian Morgenstern):

Kroklowafgi? Semememi?
Seikronto prafliplo.
Bifzi, bafzi; nulałomi ...
quasti besti bo ...

Die physische Uniformität der Textoberfläche wird vom (systematisch-wissenschaftlichen) Leser durch Einsatz verschiedener Mittel durchbrochen. Ein relevanter Sachverhalt wird etwa mit einem Markerstift hervorgehoben oder unterstrichen, Passagen werden mit Schlangenlinien im Randbereich deutlich gemacht, Zusammenfassungen, Stichworte werden angemerkt, Querverweise eingefügt, verschiedenfarbige Lesezeichen werden eingelegt - um nur ein paar der konventionellen Techniken zur "Bewältigung" eines Textes zu erwähnen.

Dient die Analyse des Textes gar der Gewinnung neuer Erkenntnisse, der genau-

en Deskription der Inhalte, der Gestaltung von Theorien, kommen weitere Techniken zur Anwendung, die sowohl mnemotechnisch zu begründen, letztlich aber auch der Kommunikation der Forscher untereinander, auch in Form von Publikationen, dienen: Mit Diagrammen, Tabellen und Grafiken werden Zusammenhänge deutlich gemacht, die der sequentielle Text eher verbirgt oder die sich im Verlauf der Interpretation aus diesem (und dem Horizont des Interpretieren) ergaben: Ein Diagramm sagt (oft) mehr als tausend Worte.

3 Netzwerke

Eine weit verbreitete Technik zur Darstellung von Zusammenhängen sind *Netzwerke*, im einfachsten Falle in Form handgezeichneter Skizzen. Netzwerke, auch Graphen genannt, bestehen im Wesentlichen aus *Knoten* und die sie verbindenden *Kanten*. Die Knoten können bspw. Textpassagen, Begriffe und Stichwörter sein. Die Kanten sind benannt oder unbenannt, oft mit Pfeilen ausgestattet und stellen etwa einen zeitlichen Ablauf (etwa im Flußdiagramm), einen Wirkungszusammenhang oder eine statische Beziehung zwischen Texten und Konzepten dar. Als abstrakte Strukturen sind Graphen Teilbereich der diskreten Mathematik mit vielen, über das rein optische hinausgehenden, nützlichen strukturellen und anderen formalen Eigenschaften, die sie als Knetmasse zur Modellierung realer Phänomene prädestinieren (Norris, 1985).

3.1 Anwendungsgebiete

Gegenüber der linearen, eindimensionalen Darstellungsform "Text" bieten Netzwerke den Vorteil, Zusammenhänge auch räumlich angemessen zu repräsentieren. Dies gilt nicht nur für die fast völlig isomorphe Modellierung von U-Bahn-Streckennetzen (oft entspricht sogar das Verhältnis der Länge der Kanten zwischen zwei Bahnhöfen der Realität), sondern auch für die Repräsentation abstrakter Sachverhalte, wie etwa einer Argumentationskette, einer komplexen Proposition, eines Entscheidungsverfahrens in Form eines Flußdiagramms oder den Sympathiebeziehungen innerhalb einer Arbeitsgruppe mittels Soziogrammen. Die "Räumlichkeit" von Netzwerkdiagrammen beschränkt sich in der Regel auf zwei Dimensionen.

Netzwerkanwendung	Knoten	Kanten
U-Bahn-Plan	Bahnhöfe	Strecken
Entscheidungsbaum	Ja-Nein-Fragen	Entscheidungen
Platinenentwurf	Bauteile	Leiterbahnen
Soziogramm	Personen	Sympathie

Argumentation	Positionen, Argumente	Gewichtungen
Mindmaps	Aspekte, Themen	Assoziationen
Gedächtnis	Inhalte	Assoziationen
Hypertext	Text	Querverweise
Thesaurus	Begriffe	Relationen
Internet	Personen, Rechner	Anrufe, Leitungen

Neue Arbeitstechniken nutzen Netzwerke oder knotenartig reduzierte Informationshappen: Die mit Karten arbeitende Metaplan-Technik etwa oder das kreative Schreiben mit Unterstützung durch Netzwerk-Skizzen (Mindmaps, v. Werder, 1992, Buzan, 1984).

Arbeitswelten selbst sind vernetzt, von streng hierarchischen Organisationsformen (Mafia, Staat) bis zu heterarchischen, mitunter chaotischen Formen der wissenschaftlichen Kommunikation über weltweite elektronische Netzwerke (etwa das Internet, siehe Broad, 1993). Die Simulation komplexer Systeme, wie etwa das Weltmodell des Club of Rome (Eigen & Winkler, 1975) oder eine Modellwelt zur Untersuchung menschlichen Entscheidungsverhaltens (Dörner, 1989) basiert auf miteinander vernetzten Einflußgrößen.

Retrieval-, Dokumentations- und Publishingsysteme verwenden zunehmend mehr oder wenige explizite Netzwerkrepräsentationen. Hypertextsysteme, Hypermedia-systeme (Kuhlen, 1991; Bogaschewsky, 1992; Barrett, 1989; Parsaye et.al., 1989; Rada, 1991; Brassel, 1992) erlauben die technische Verkopplung von Textteilen, aber auch die Definition einer prozeduralen Semantik: beim Anklicken einer Textpassage wird eine andere Passage angezeigt oder die Bandstelle (Video, Audio) des Interviewtextes vorgespielt. Wissensakquisitionssysteme unterstützen die Wissensdarstellung in strukturierter und formalisierter Form, als semantische Netzwerke, Frames, Scripts oder logikbasierten Kalkülen (Anjewierden 1987). Auch die Vernetzung der Menschen untereinander wird technisch unterstützt: E-mail Systeme fördern die Kommunikation am Arbeitsplatz, Groupware unterstützt komplexe arbeitsweltliche Diskurse, bspw. die Terminplanung oder das gemeinsame Bearbeiten komplexer Dokumente (Ko-Autorenkonzept in ATLAS/ti). Systeme zur Unterstützung der Textinterpretation und der qualitativen Datenanalyse erlauben den Aufbau von auf textuellem Ausgangsmaterial basierenden Theorien und die Nutzung dieser Strukturen für das Retrieval (Muhr, 1991 und in diesem Band; Richards & Richards, 1994; Dey, 1992).

3.2 Netzwerke: "Mehrwert" für die Textinterpretation?

Unschwer festzustellen: Das Thema "Netzwerke" hat - wie die Netzwerke selbst - die Tendenz zur Ausbreitung. Wir wollen uns hier jedoch auf deren Nutzung für die textinterpretation beschränken. Netzwerke können im Rahmen der Textinter-

pretation verschiedene Funktionen erfüllen. Zunächst läßt sich fragen, ob und in welcher Weise das untersuchte Material und die im Verlauf der Interpretation herausgearbeiteten Sachverhalte und Hypothesen durch Netzwerke an Aussagekraft oder sonstigen Eigenschaften gegenüber der linearen Textform gewinnen. Mit Kuhlen (1991), der dies für die Nutzenabschätzung von Hypertextsystemen so prägnant formuliert, kann man fragen, welchen "Mehrwert" Netzwerke für die Textinterpretation besitzen. Schließlich steht am Anfang des Prozesses der Textinterpretation ein - zu interpretierender - linearer Primärtext und das Resultat ist wiederum ein sequentieller, publikationsfähiger Text.

Für die Verwendung von Netzwerken im Rahmen der Textinterpretation sprechen verschiedene Gründe, die sich - neben anderen - einerseits aus Annahmen über unseren Denk- und Wahrnehmungsapparat, andererseits aus informatischen Erkenntnissen herleiten.

3.2.1 Netzwerke & Kognition

Die kognitive Position (Anderson, 1989) geht von einem Netzwerkmodell menschlichen Wissens und Denkens aus. Aus kognitionswissenschaftlicher Sicht sind Gedächtnisinhalte nicht sequentiell in durchnummerierten Schubläden abgelegt, sondern in komplexer Weise miteinander vernetzt. In unterschiedlichen Auffassungen spielen sowohl die Knoten als auch die Verbindungen zwischen den Bedeutungs"Knoten" eine unterschiedliche Rolle. Rein *assoziative* Gedächtnismodelle gehen von durch unterschiedliche Stärke verbundenen Gedächtnisinhalten aus¹. In *semantischen* Netzwerken sind die Verbindungen dagegen symbolisch bewertet, z.B., die Verbindung "ist-Unterklasse-von" (isa) zwischen zwei Konzepten. Erinnerung wird so als Prozess der Aktivierungs-Ausbreitung ("spreading activation") über benachbarte Knoten bzw. als Ergebnis logischer Ableitungen innerhalb des Gedächtnisnetzwerkes und nicht als Lesen von Speicherstellen (wie wüßte man auch, welche Speicherstellen adressiert werden müßten?) erklärt. Diese beiden Auffassungen über die Natur menschlicher Kognition finden sich übrigens auch in der Kontroverse über die Künstliche Intelligenz wieder (Dreyfus, 1987, 1990).

Die Darstellung von Bedeutungen als Netzwerk-Diagramme bringt wegen dieser Nähe zu unserem kognitiven System Vorteile: Wissen wird kontextuell dargestellt und kann so Zusammenhänge sichtbar machen, die in rein textueller Form eher verdeckt bleiben. Strukturierte Wissensdarstellungen können im Rahmen der Textinterpretation die Funktion einer intuitionsfördernden, intermediären Repräsentation einnehmen, die nicht Selbstzweck ist, sondern letztlich die Produktion neuen Textwissens begleitet:

¹ Als neuronale Netze erleben assoziativ-subsymbolische Netzwerke eine Renaissance

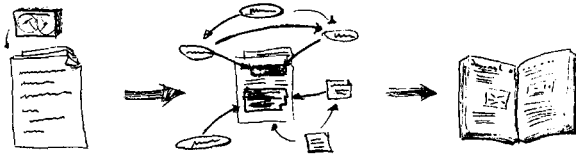


Abb. 1: Text - Struktur - Text

Diese - synergetische - Eigenschaft von Netzwerken, Sachverhalte direkt und in ihrem jeweiligen Zusammenhang darzustellen, wird in den bereits erwähnten kreativen Schreibtrainings (Mindmaps) und Moderationstechniken (Metaplan) genutzt.

Kognitive Ökonomie ist eine weitere Eigenschaft semantischer Netzwerke. Merkmale, die für Konzepte gelten, die in einer höheren Generalisierungshierarchie liegen, müssen nicht explizit bei jedem Unterkonzept angeführt werden. Der Fakt, daß "Lebewesen" einen Stoffwechsel haben, wird so nur einmal gespeichert und an die spezifischeren Konzepte "vererbt", wenn sie, bspw. in einem Wahrnehmungs- oder Erinnerungsprozess benötigt werden. Durch "spreading activation" wird bspw. vom Knoten "Fisch" das Netzwerk so lange durchforstet (in diesem Falle über die transitive "ist-Unterbegriff-von" Beziehung) bis das Merkmal "Stoffwechsel" gefunden ist.

Eine Speicherung von Erinnerungen in textueller Form wäre dagegen in hochgradig ineffizient. Nach Buzan (1991) sind 90% der Wörter eines Textes für die Erinnerung irrelevant.

3.2.2 Netzwerk-Mehrwert aus Informatik-Sicht

Die kognitive und die informatische Position ist nicht scharf voneinander zu trennen, wenn von den Erkenntniszwecken abgesehen wird, da beide Disziplinen - besonders deutlich bei der Terminologie - von der jeweils anderen beeinflusst wurden.

In der Informatik werden Netzwerkformalismen u.a. für die Wissensrepräsentation eingesetzt. Informelle Eigenschaften, wie Prägnanz, werden zum einen unter Rückgriff auf die kognitive Position begründet. Die Darstellungseigenschaften von Netzwerken sind wichtig für die bei der Wissensakquisition in der Phase der Wissenserhebung stattfindende Kommunikation zwischen Wissensingenieur und Fachexperten. Anhand der Diagramme können Prozesse und Wissensstrukturen transparent gemacht werden.

Als alternative Darstellung zu linearem Text sollten semantische Netzwerke die für eine technische Interpretation geeignete Repräsentation textuellen Wissens ermöglichen (Quillian, 1968). Diese Hoffnung hat sich trotz verschiedener Wei-

terentwicklungen des ursprünglichen Ansatzes (z.B. in der Wissensrepräsentationssprache KL-ONE, Brachmann & Schmolze, 1986; Wielinga et. al., 1992) bis heute nicht erfüllt. Verschiedene Modelle und Implementierungen auf Basis von Netzwerkformalismen haben jedoch durchaus zu neuen - auch kognitionspsychologischen - Erkenntnissen und technischen Anwendungen geführt (Sowa, 1988). Das schon erwähnte Prinzip der Merkmalsvererbung im Sinne einer kognitiven Ökonomie dient bei der Umsetzung in technische Systeme der effektiven Nutzung von knappen Computerressourcen und war ein wichtiger Grund für die Verwendung von Netzwerkformalismen bei der maschinellen Repräsentation von Wissen. Als Formalismus für die Darstellung von (statischem) Wissen in *Expertensystemen* kann zudem auf die Rückführbarkeit semantischer Netze auf das Prädikatenkalkül 1. Stufe verwiesen werden. Dessen formale Eigenschaft der Entscheidbarkeit erlaubt Inferenzen auf dem vernetzten Datenmaterial durch rein syntaktische Mittel (also auch durch Computer). Das in der folgenden Abbildung dargestellte Netzwerk

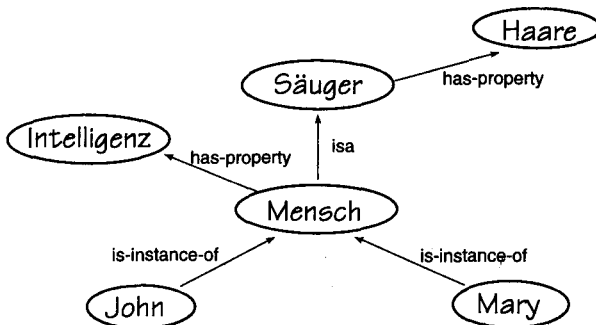


Abb. 2: Semantisches Mininetzwerk

reduziert sich im Prädikatenkalkül auf eine Menge zweistelliger Prädikate:

- isa (Mensch, Säuger).
- has-property (Säuger, Haare).
- has-property (Mensch, Intelligenz).
- is-instance-of (John, Mensch).
- is-instance-of (Mary, Mensch)².

Das Prädikatenkalkül erlaubt die Ableitung ("Vererbung") des - nicht explizit beschriebenen! - Fakts, daß sowohl John als auch Mary intelligent sind und Haare haben. Dazu werden allerdings Eigenschaften der beteiligten Relationen und Vererbungsregeln (z.B. Transitivität) im Kalkül zusätzlich beschrieben, die aus

² isa = ist Unterklasse bzw. Teilmenge von, has-property = hat Eigenschaft, is-instance-of = ist Ausprägung, bzw. Element von

der Netzwerkdarstellung allein nicht deutlich werden. Das obige Beispiel ist leider nicht so ganz trivial, da die Ableitungen zwar als Vorannahme gelten können, aber keineswegs regelmäßig zutreffen müssen. Wie stellt man bspw. dar, daß John Glatzenträger ist? Solche Probleme sind Gegenstand von auf nicht-monotonen Logiken basierenden Ansätzen. Anzumerken ist, daß das Prädikatenkalkül erheblich ausdrucksstärker im Sinne vollständiger und korrekter Beschreibungen ist. Bspw. lassen sich im Netzwerk Quantoren (etwa "alle", "einige", "genau ein") nur mühsam ausdrücken und führen zu zwar korrekteren und vollständigeren, dafür aber deutlich weniger prägnanten Darstellungen. Bei Netzwerk-Anwendungen, die deduktive Mechanismen einschließen, kommt man um eine formal-semantische Fundierung, etwa im Prädikatenkalkül erster Ordnung nicht herum (zu einer denotationalen und modell-theoretischen Semantik von KL-ONE siehe auch v.Luck & Owsnicki-Klewe, 1987).

Der Begriff des Netzwerks hat in der Informatik zusätzlich die Konnotation des Prozeduralen. Mit komplizierten Algorithmen steigt allerdings der Aufwand für die Entwicklung, Wartung und den Betrieb von Systemen, die auf Netzwerken basieren. So wurden bspw. die auf einem Netzwerkmodell basierenden Datenbanksysteme (etwa nach dem CODASYL-Modell), die den Anfragenden quasi zu einem "Piloten" machten, der per eingebetteter Programmanweisungen durch das Netz navigiert, durch die relationalen, deskriptiven Datenbankmodelle abgelöst. Mit den objektorientierten Sprachen und Datenbanksystemen erleben netzartige Ansätze eine Renaissance in konventionellen (nicht: Künstliche Intelligenz) Anwendungen, allerdings mit teilweise erheblich verbesserten Modellen und Verfahren. (Breutmann & Burkhardt, 1992; Budde et.al., 1989).

Auch Hypermediasysteme mit ihren Knoten, Kanten und Pfaden basieren auf Netzwerkmodellen. Die Benutzeroberflächen sind hierbei optimal für die Navigation, das "Stöbern" (Browsing) in den vernetzten Datenbeständen ausgelegt. Inwieweit diese Systeme sich auch für große Datenmengen als erfolgreich erweisen oder etwa im Verbund mit anderen Retrievaltechniken optimiert werden können, bleibt vielversprechendes Forschungsthema (Fuhr, 1992).

Der Mehrwert von Netzwerken für die Textinterpretation erscheint also auf verschiedenen Ebenen:

Als "cognitive aid" zur visuellen Unterstützung unseres ohnehin netzbasierten kognitiven Apparats, als Kommunikationsmedium und als Mittel zur Strukturierung und Speicherung von Texten und Bedeutungen für den Aufbau von Theorien und für das Information-Retrieval.

4 Netzwerke für die Textinterpretation

Was kann nun bei der Textinterpretation mit Hilfe von Netzwerken dargestellt werden?

1. Die Textoberfläche selbst wird als Graph dargestellt. Bei entsprechender syntaktischer Qualität des Ausgangstextes können solche Syntaxbäume mit Hilfe geeigneter Parser (Programme, die den Text in die Konstituenten zerlegen) auch automatisch generiert werden. Diese Form der Anwendung wird wohl vor allem linguistischen Fragestellungen vorbehalten bleiben, wird aber auch bei der Realisierung automatischer Übersetzungssysteme angewendet.
2. Die Inhaltsebene, d.h., die im Texte selbst enthaltenen Konzepte werden in Form von Netzwerken dargestellt (wie in Abb. 11). Netzwerke sind hierbei Mittel einer genaueren Deskription, mit der auch Mehrdeutigkeiten des Textes reduziert werden können (siehe auch das Kranich-Beispiel auf Seite 301).
3. Aus dem Text per Interpretation abgeleitete oder hinzukommende Konzepte werden vernetzt. Hierbei wird die Ebene reiner Deskription verlassen und man kommt zu Netzwerken, die Teile einer Theorie repräsentieren oder eine reduzierte Deskription des Textes darstellen (semantische Netzwerke, Thesauri).
4. Textpassagen werden untereinander zu Hypertexten vernetzt. Bsp.: "Wir wollen nichts als Frieden!" steht-in-Widerspruch-zu "Die anderen haben immer Unrecht".

Diese verschiedenen Anwendungen von Netzwerken zur Repräsentation von Textwissen korrespondieren zu der Wissensrepräsentations-Typologie nach Frawley (1988): dem propositionalen System, dem rhetorischen System und dem konzeptuellen System.

Welcher Art sind nun die für die Vernetzung zu verwendenden Relationen? Womit werden jeweils zwei Begriffe, Konzepte, Textstellen verbunden? Dies ist natürlich nicht unabhängig von der obigen Kategorisierung. Wir beziehen uns deshalb im folgenden auf propositionale und konzeptuelle Repräsentationsformen.

Die Entscheidung für die "richtigen" Relationen ist gleichbedeutend mit der Entscheidung für einen bestimmten Blickwinkel, aus dem heraus das vorliegende Material untersucht wird. Sie bilden die "erkenntnisleitenden Primitive" (epistemological primitives, Brachman & Schmolze, 1986), die die Fragen an das Material - den Text und die aus diesem abgeleiteten Konzepte - lenken.

Für semantische Netzwerken haben sich Relationen wie "is-a" (Unterbegriffs-Overbegriffs-Beziehung, auch als Teilmengen-Beziehung zu betrachten: "Frau is-a

Mensch"), "ist-Teil-von" (Teil-Ganzes-Beziehung) oder "ist-Instanz-von" (Element-Klassen-Beziehung: "Mary ist-Instanz-von Frau") etabliert. Diese Relationen haben statisch-definitiven Charakter und werden in der Regel für die Beschreibung von relativ beständigem Wissen verwendet. Strukturierte Termbestände, wie sie beim Information-Retrieval zur Anwendung kommen, sind meist durch wenige Relationen obiger Art vernetzt. Solche *Thesauri* können - von der Struktur, nicht vom Zweck her - als Spezialfall semantischer Netzwerke betrachtet werden. (Burkart, 1988).

Die terminologischen Relationen (*isa*, *is-part-of*, *is-synonym-of*) sind zur Ordnung von Dokumentendeskriptoren, für Zwecke normenorientierter Informationsspeicherung und einem darauf basierenden Retrieval gut geeignet, jedoch als Ausdrucksmittel zur Repräsentation textuellen Wissens oft zu "mager". Manche Relationen sind schon für sich nicht unkompliziert. So gibt es Studien zur der Partonomie-Relation (Teil-von), die eine unglaubliche Vielfalt der Interpretationen belegen. So läßt sich darunter u.a. das Verhältnis von Kopf und Körper, Wein und Flasche, Monat zu Jahr, Fahrer zu Auto, Motor zu Auto, Darsteller zu Film, etc. subsumieren (Chaffin & Herrmann, 1988).

Die in ATLAS/ti zur Verfügung gestellten Standardrelationen beinhalten eine willkürliche Mixtur aus terminologischen und sonstigen semantischen Relationen (*isa*, *is-part-of*, *is-cause-of*, *is-property-of*, etc.). Interessanterweise konnten wir beobachten, daß Anwender die vorgegebenen Relationen verwendeten, ohne die Möglichkeit zu nutzen, angemessenere Relationen selbst zu definieren. Der vorliegende Artikel soll auch gerade dazu ermuntern.

4.1 Modellbasierte Netzwerke

In Abgrenzung zu dem terminologischen, semantischen Wissen kann in Anlehnung an die in Lindsay & Norman (1981) beschriebene Unterscheidung von semantischem und episodischem Gedächtnis bei der Darstellung von Ereignissen von episodischem Wissen gesprochen werden. Eine daran angelehnte Unterscheidung findet sich in der Wissensakquisitionssprache KL-ONE als T-BOX (terminologisches Wissen) und A-BOX (assertatorisches Wissen, Faktenwissen) zur Trennung der Wissensarten.

Während sich statisch-semantisches Wissen in relativ einfachen, hierarchischen Diagrammen darstellen läßt (thesaurusartigen Ober-Unterbegriffs-Hierarchien), ist episodisches Wissen komplexer strukturiert. Die Relationen, die zur Vernetzung von Objekten und Ereignissen zur Darstellung episodischen Wissens verwendet werden, sind vielfältiger und lassen sich meist einem erkenntnisleitenden "Modell" zuordnen. Das modellbasierte Vorgehen beinhaltet die Trennung in zwei Ebenen: die des *anwendungsspezifischen Wissens* ("domain-level-knowledge")

und eine diese organisierende Ebene des Meta-Wissens ("meta-level-knowledge"). Bei der Wissensakquisition wird ein solches schichtenorientiertes Vorgehen u.a. im Rahmen der KADS-Methodologie³ verfolgt (Wielinga & Breuker, 1986; Wielinga, van de Velde, Schreiber, 1992). Wesentlich für den Einsatz modellbasierter Verfahren für die Textinterpretation und die Theorieentwicklung ist eine Änderung der Arbeitsweise: Während beim "Aufbrechen" der Daten und der Extraktion von Begriffen aus den Texten eine "bottom-up"-Strategie verfolgt wird, erfolgt die Strukturierung der Begriffe zu theoretischen Netzwerken "top-down" - gesteuert durch das im Modell realisierte methodische Konzept. Der Wechsel zwischen diesen beiden Bewegungsrichtungen ist wesentlicher Bestandteil des Grounded-Theory-Verfahrens (s.u.).

4.1.1 Beispiel: Handlungsmodell

Strukturelle Modelle, etwa das Handlungsmodell, sind im Zusammenhang mit linguistischen Untersuchungen auch als Kasusgrammatiken (Filmore, 1968) bekannt. Der Kasus bezeichnet die Rolle, die bspw. eine Nominalphrase in Bezug auf die durch das Verb bezeichnete Aktivität spielt. Ähnlichkeit besteht auch zu den propositionalen Netzwerken (Anderson, 1989, v. Dijk & Kintsch, 1983) zur Repräsentation von Textwissen.

Die Wahl eines Modells ("Paradigma") ist geprägt vom spezifischen Erkenntnisinteresse. Das in Abb. 3 dargestellte Handlungsmodellschema stellt die Handlung in den Mittelpunkt, ein aktorsorientierter Ansatz fokussiert dagegen auf die Handelnden einer Situation.

³

KADS = Knowledge Acquisition and Document Structuring

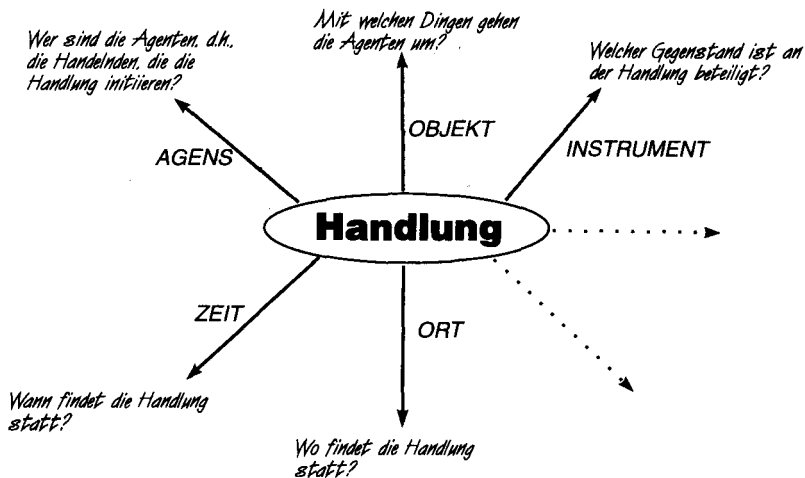


Abb. 3 Handlungsmodell nach Lindsay/Norman 1981.

Ein Modell stellt zunächst eine Menge von Relationen (Rollen) zur Verfügung. Im Handlungsmodell sind das: *agens*, *objekt*, *instrument*, *ort*, *zeit*, usw. Mit diesen Relationen werden dem zentralen Objekt, hier der Handlung, die handlungsbeschreibenden Konzepte zugeordnet. Die Relationen weisen den beteiligten Konzepten gewissermaßen ihre Rolle innerhalb der so modellierten Aussage, bzw. Proposition, zu. Aus diesem Grund können die Relationen auch als *Rollen* aufgefaßt werden. Ein Unterschied zu den terminologischen Netzwerken ist, daß die "Relation", die alle Bestandteile zusammenhält, eben die Handlung, selbst als Knoten und nicht als (mehrstellige) Relation erscheint. Wir haben es nun also mit unterschiedlichen Knotentypen zu tun.

Die folgende Abbildung zeigt drei als Netzwerk dargestellte Interpretationen eines mehrdeutigen Satzes unter Anwendung des Handlungsmodells. Man beachte, daß die Handlungen mit Suffixen versehen sind. Damit wird deutlich gemacht, daß es sich um konkrete Ausprägungen, Instanzen, einer Handlungsart handelt. Es wird nicht SEHEN als solches dargestellt, sondern ein konkretes Sehens-Ereignis, eben episodisches Wissen.

Müllers sahen die Kraniche, als sie über die Alpen flogen

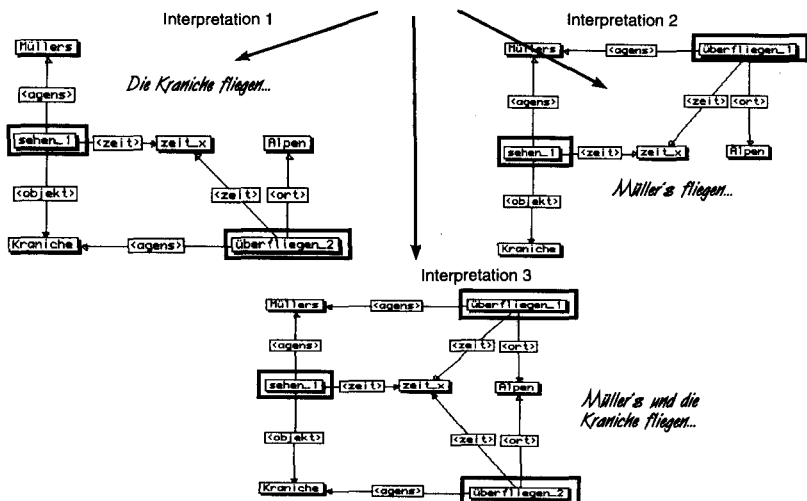


Abb. 4: Propositionale Netzwerke disambiguieren einen mehrdeutigen Satz - falls man sich entscheiden kann. Die Netzwerke wurden mit ATLAS/ti erstellt.

Die Ähnlichkeit der Relationen/Rollen und deren Belegung zu den "Slots & Fillers" der Frame-Theorie wird deutlich in linearer Notation⁴. Die Interpretation 1 aus Abb. 4 beschreibt zwei gleichzeitig stattfindende Handlungen, eine Sehens- und eine Überfliegens-Handlung:

SEHEN_1

agens: Müllers
objekt: Kraniche
zeit: zeit_x

ÜBERFLIEGEN_2

agens: Kraniche
ort: Alpen
zeit: zeit_x

Mit den Rollen steht dem Interpreten, der seinen Texten mit Netzwerken zu Leibe rückt, ein Instrumentarium zur Ordnung der Konzepte zur Verfügung, die sich im Verlauf der Analyse der Texte ergeben haben. Dabei sind solche Modelle nicht nur für die Beschreibung und Einordnung, sondern auch schon bei der Erhebung und Interpretation des Datenmaterials nützlich: wichtige Aspekte können systematisch erhoben werden, fehlendes Material kann gezielt nachgefordert werden. In Abb. 3 sind die bei Anwendung des Handlungsmodells relevanten Fragen aufgeführt. Wie die einzelnen Konstituenten in der manifesten Oberfläche des analysierten erscheinen, kann dabei durchaus recht vielfältig sein. Die folgende Abbildung zeigt drei verschiedenen Realisierungen eines Bedeutungsnetzes:

⁴ Mit ATLAS/ti kann eine solche lineare textuelle Darstellung automatisch generiert werden.

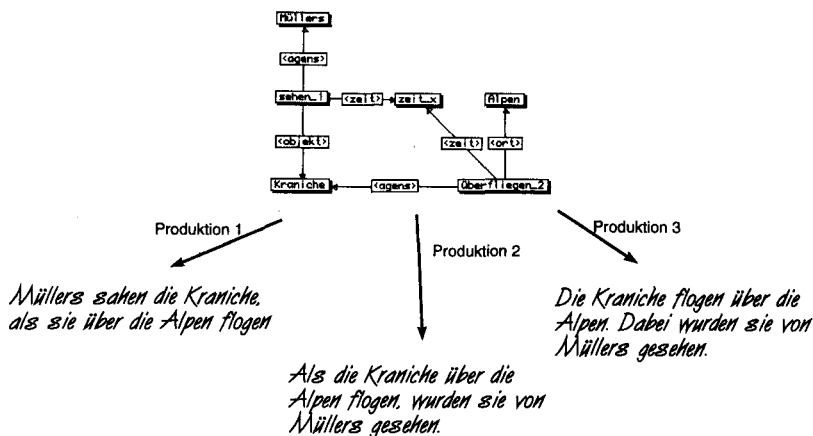


Abb. 5: Einem propositionalem Netzwerk entsprechen verschiedene Textmanifestationen.

Eine wichtige Eigenschaft solcher "propositionalen" Netzwerke ist die Einführung von Ereignisknoten, wie das Sehen oder das Überfliegen. Diese Knoten integrieren die relevanten Konzepte, ähnlich wie dies ein sprachlicher Satz tut. Damit ist es möglich, mehrstellige Relationen auf Bündel zweistelliger Relationen zu reduzieren - notwendige Voraussetzung für eine Netzwerkdarstellung. Dies entspricht auch der Transformation von Prädikaten in eine objektorientierte Repräsentation über Skolemisierung (UNITS, siehe auch Nilsson, 1982, S. 362ff).

4.1.2 Beispiel: Paradigmenmodell nach Strauss

Ein Erkenntnismodell, das für die Analyse von Texten unter sozialwissenschaftlicher Perspektive eingesetzt wird, ist die Grounded Theory-Methode (GT) von Glaser & Strauss (1967). Eine Beschreibung und Anwendungsbeispiele finden sich in verschiedenen Beiträgen in diesem Band (Böhm, Engelmeier).

Im folgenden Kapitel wird näher auf die im Verlauf der GT sich bildenden Strukturen eingegangen. Zu Beginn soll kurz auf die spezifische Form des "Vernetzens" im Rahmen des axialen Kodierens (s.u.) und die strukturelle Ähnlichkeit zu anderen Modellen, wie etwa dem oben skizzierten Handlungsmodell hingewiesen werden:

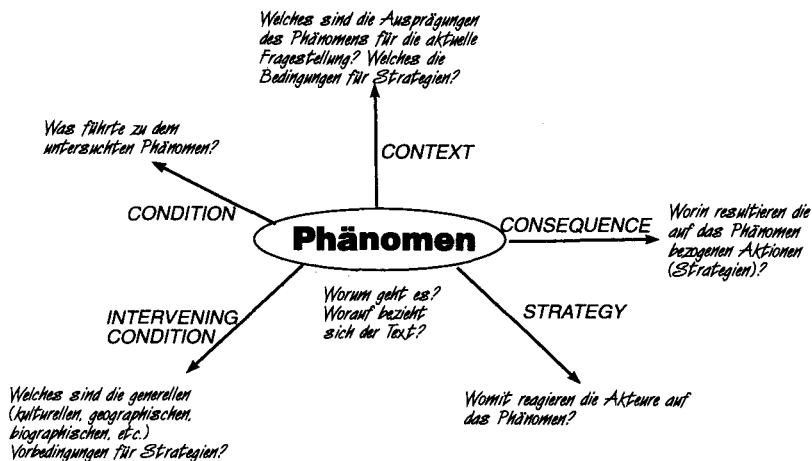


Abb. 6: Paradigmenmodell nach Strauss & Corbin.

Auch hier haben die "Rollen" erkenntnisleitende Funktion, die die Fragen an das Datenmaterial vorgeben. Im Zentrum der Betrachtung steht hier das "Phänomen", entsprechend der Handlung im Handlungsmodell.

So wie hier exemplarisch für das handlungs- und phänomenorientierte Vorgehen lassen sich weitere Modelle konstruieren, die eine gegenstandsangemessene Analyse fördern: etwa für den Bereich der medizinischen Diagnose mit den Rollen (bzw. Meta-Kategorien): Diagnose, Verdacht, Symptom, Syndrom, Medikation und Therapie.

Die "Grounded Theory" ist eine Methode der Qualitativen Sozialforschung, die aus einer Reihe von Regeln besteht, mit denen eine in den Daten begründete Theorie gewonnen werden kann. Sie wurde von B. Glaser und A. L. Strauss (1967) mit "The Discovery of Grounded Theory" in einer ersten Zusammenfassung vorgestellt. Wir beziehen uns hauptsächlich auf das Buch "Basics of Qualitative Research" von A. L. Strauss und J. Corbin (1990).

Wir wollen hier die Methode in aller Kürze beschreiben und die dabei entstehenden Datenstrukturen mit Diagrammen veranschaulichen.

Am Anfang der Datenanalyse steht das Finden von *Konzepten*. Konzepte sind aus den Basisdaten abgeleitete oder entnommene ("in-vivo-Codes") Namen oder Begriffe.

Die gefundenen Konzepte werden dann zu *Kategorien* und Unterkategorien zusammengefaßt. Eine Gruppe von Konzepten wird jeweils einer Kategorie zugeordnet. Dabei trägt diese Kategorie einen abstrakteren Namen als die Konzepte, die sie beinhaltet.

Jede Kategorie wird außerdem durch eine Menge von *Eigenschaften* näher beschrieben. Diesen Eigenschaften sind wiederum *Dimensionen* zugeordnet. Diese sind Wertebereiche der Eigenschaften, wobei jede Eigenschaft zusätzlich Unter-eigenschaften besitzen kann.

Die erste Phase der Analyse, zu der diese Strukturierungen gehört, wird "offenes Kodieren" genannt.

Konzepte besitzen zu allen Eigenschaften der übergeordneten Kategorie einen konkreten Wert. Eine Kategorie mit ihren Eigenschaften stellt gewissermaßen ein Modell ihrer subsumierten Konzepte dar.

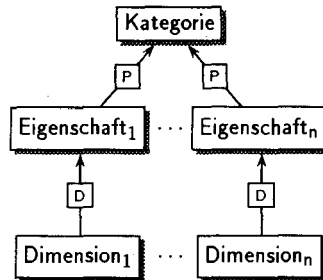


Abb. 7: Kategorie mit Eigenschaften und deren Dimensionen. "P" und "D" stehen für "Property" und "Dimension".

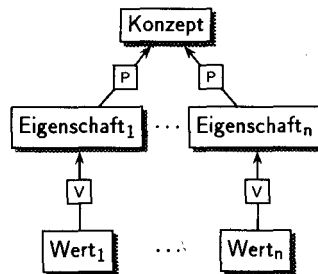


Abb. 8: Konzept mit der übergeordneten Kategorien entsprechenden Eigenschaften und Werten. "V" steht hier für "Value".

In der zweiten Phase, dem "axialen Kodieren", werden die gewonnenen Kategorien untereinander nach einem vorgegebenen Modell, dem sogenannten *Paradigmenmodell*, in Beziehung gesetzt.

Dieses Modell beinhaltet folgende Rollen:

- (A) Ursächliche Bedingungen → (B) Phänomen →
- (C) Kontext → (D) Intervenierende Bedingungen →
- (E) Aktion/Interaktive Strategien →
- (F) Konsequenzen.

Diese Rollen hängen folgendermaßen zusammen: die *ursächlichen Bedingungen* lösen das *Phänomen* (mit seinem spezifischen *Kontext*) aus, Anwendung von *Aktion/Interaktiven Strategien* in Bezug auf ein *Phänomen* (in einem spezifischen *Kontext*) führt zu *Konsequenzen*.

Einem spezifischen Phänomen können also jeweils mehrere *Nachbarkategorien* als "ursächliche Bedingung" etc. zugeordnet sein.

Eine Kategorie, die in einem Paradigmenmodell "Phänomen" ist, kann innerhalb einer anderen Modellausprägung die "ursächliche Bedingung" oder "Konsequenz" darstellen.

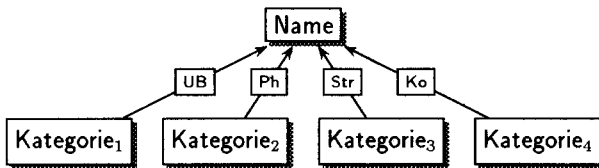


Abb. 9 Paradigmenmodell-Schema.

Abb. 9 zeigt ein Paradigmenmodell mit "ursächlichen Bedingungen" (UB), dem "Phänomen" (PH), "Strategien" (Str) und "Konsequenzen". "Kontext" und "intervenierende Bedingungen" sind Eigenschaften der Kategorien zu "Phänomen" und "Strategien".

Die verschiedenen, mit Kategorien versehenen, "Paradigmenmodelle" werden dann durch entsprechende Belege (Konzepte) in den Daten validiert.

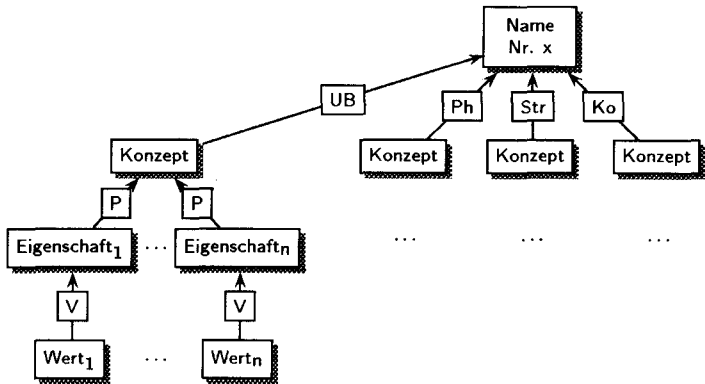


Abb. 10: Beleg zu einem Paradigmenmodell.

Das Paradigmenmodell bildet also ein Schema, mit dessen Hilfe aus Kategorien Schemata für die beteiligten Konzepte gewonnen werden, gewissermaßen also ein *Meta-Schema*.

Corbin & Strauss geben ein Beispiel eines ausgearbeiteten Modells:

ursächliche Bedingung ein Beinbruch	→	Phänomen Schmerz
<i>Eigenschaften eines Beinbruchs</i>		<i>Dimensionen des Schmerzes</i>
Mehrfachfraktur		Intensität stark
komplizierter Bruch		Dauer andauernd
offene Fraktur		Position Unterschenkel
vor zwei Stunden gebrochen		Verlaufskurve früh
Sturz im Wald		Hilfsaussicht lange Wartezeit
		Potential für hoch
		Konsequenzen

Kontext der Schmerzbehandlung

Unter der Bedingung, daß der Schmerz:
intensiv, andauernd, im Unterschenkel, früh in der Verlaufskurve
vorhanden ist und:
die Wartezeit auf Hilfe lang, das Potential für Konsequenzen hoch
ist, dann sind:

Strategien der Schmerzbehandlung
 das Bein schienen
 den Rettungsdienst benachrichtigen
 die Person warmhalten

Intervenierende Bedingungen
 fehlendes Training in erster Hilfe
 das Fehlen einer Decke
 ein weiter Weg, Hilfe zu holen

Sie legen dabei auf die Untersuchung der *Eigenschaften* des Phänomens (diese bilden den "Kontext") besonderen Wert.

In der dritten Phase, dem "selektiven Kodieren", wird aus den so gewonnenen (und in den Basisdaten verankerten) Paradigmenmodellen eine Theorie entwickelt. Die Aktivitäten der zweiten und dritten Phase gehen dabei fließend ineinander über.

5 Formalisierung struktureller Aspekte der Grounded Theory

Es folgt ein Versuch einer formalen Beschreibung der im Verlauf des Kodierens entstehenden Strukturen mit Hilfe sog. Signaturen (siehe auch Seibt, 1993). Sie werden von Informatikern für die Beschreibung von Datenstrukturen verwendet. Hiermit lassen sich präzise und implementierungsunabhängig die für eine eventuell maschinelle Bearbeitung notwendigen Aspekte beschreiben.

Bereiche, die sich anhand der Beschreibungen nicht präzise spezifizieren lassen, sind die Basisdaten (werden so benannt) und die durch die Dimensionen repräsentierten Wertebereiche (werden mit 'Data' bezeichnet). Diese werden als gegeben vorausgesetzt, ebenso der Datentyp 'Menge' sowie der Typ 'Bool' für Wahrheitswerte und der Typ 'Name' für die Namen von Konzepten und Kategorien.

signatur Zitat =

Bewirkt eine Markierung in den Daten

sorts:

Zitat, Basisdaten

opns:

$\langle \text{zitat} : \text{Basisdaten} \rightarrow \text{Zitat} \rangle$

signatur Konzept =

sorts:

Konzept, Name, Zitat, Kategorie

opns:

$Name \times Zitat \times \text{übergeordnete Kategorie} \times \text{Bewertungsfunktion}$

$\langle \text{konzept: Name} \times \text{Zitat} \times \text{Kategorie} \times (\text{Kategorie} \rightarrow \text{Set}[\text{Eigenschaft} \times \text{Data}]) \rightarrow \text{Konzept} \rangle$

Die Bewertungsfunktion der Konzepte ordnet jeder Eigenschaft der übergeordneten Kategorie einen konkreten Wert zu.

signatur Kategorie =

ordnet einer Menge von Konzepten (eine Menge von) Eigenschaften zu

sorts:

Kategorie, Name, Menge[Data], Konzept, Eigenschaft

opns:

$\langle \text{kategorie: Name} \times \text{Menge}[\text{Konzept}] \times \text{Menge}[\text{Eigenschaft}] \rightarrow \text{Kategorie} \rangle$

signatur Eigenschaft[Data] =

Eigenschaft, Wertebereich und Untereigenschaften

sorts:

Eigenschaft[Data], Dimension[Data]

opns:

$\text{Wertebereich} \times \text{Untereigenschaften}$

$\langle \text{eigenschaft[Data]: Dimension[Data]} \times \text{Menge}[\text{Eigenschaft}] \rightarrow \text{Eigenschaft[Data]} \rangle$

signatur Dimension[Data] =

Modellierung eines Wertebereichs

sorts:

Dimension[Data], Data, Menge[Data]

opns:

$\langle \text{Dimension[Data]: Menge[Data]} \rightarrow \text{Dimension[Data]} \rangle$

value : Dimension[Data] \rightarrow Data

signatur PM =

Paradigmenmodell

sorts:

PM, Kategorie, Set[Data]

opns:

ursächl. Beding. \times *Phänomen* \times *interv. Beding.* \times *Aktion/Interaktion* \times *Konseq.*

$\langle \text{pm: Set[Kategorie]} \times \text{Kategorie} \times \text{Set[Kategorie]} \times \text{Set[Kategorie]} \times \text{Set[Kategorie]} \rightarrow \text{PM} \rangle$

Hier wird ein Paradigmenmodell auf der konzeptuellen Ebene beschrieben. Um eine konkrete Ausprägung zu erhalten, müßte für jede der Rollen eine Funktion hinzugefügt werden, die der Menge der Kategorien eine Menge der konkreten Konzepte zuordnet ($\text{Set[Kategorie]} \rightarrow \text{Set[Konzept]}$) bzw. im Fall des "Phänomens" der Kategorie ein Konzept zuordnet ($\text{Kategorie} \rightarrow \text{Konzept}$).

6 Probleme beim Einsatz von Netzwerken für die Textinterpretation

Ein Problem, sowohl bei der "heuristischen" Verwendung von Netzwerken zur Unterstützung und Anregung kreativer Denkprozesse, als auch bei der maschinellen Repräsentation von Wissen, besteht darin, Knoten und Kanten "richtig" zu benutzen (Kritik an semantischen Netzwerken: Woods, 1975). Analoges trifft auch für den Bereich der Textinterpretation zu. Bereits beim *offenen Kodieren* der Grounded Theory, vereinfacht gesagt, dem Finden und der Zuordnung von Stichwörtern zu Textpassagen, stellt sich die Frage nach der Bedeutung der Codes im Verhältnis zu den von ihnen referenzierten Textstellen. Sind Codes nichts weiter als abkürzende Namen für gruppierte Textpassagen, sind Textpassagen empirische Belege oder Definitionen für in Codes repräsentierten, theoretisch bedeutsamen Konzepten?

Wie brauchbar sind netzwerkorientierte Darstellungsformen bei längeren Texten? Schon bei der Darstellung einfacher sprachlicher Phänomene zeigt sich schnell der Flächenbedarf einfacher Netzwerke. Abb. 11 ist die grafische Repräsentation des folgenden Kurztextes:

"A great black and yellow V-2 rocket 46 feet long stood in a New Mexico desert. Empty, it weighted five tons. For fuel it carried eight tons of alcohol and liquid oxygen. Everything was ready. Scientists and generals withdrew to some distance and crouched behind earth mounds. Two red flares rose as a signal to fire the rocket."

befindliche Netzwerksicht einbezogen werden. Störende "Knoten" und die von ihnen ausgehenden "Kanten" können wahlweise ausgeblendet werden. Problematisch bei komplexen Netzwerken ist auch die Anordnung der Knoten und die Führung der Verbindungen unter Vermeidung eines unübersichtlichen "Verhaus". Intelligente Layoutverfahren zum geschickten Plazieren der Knoten können dabei eine große Hilfe sein (Seibt, 1992).

Woods (1975) hat u.a. auf das Problem der Darstellung mehrstelliger Relationen in Netzwerken aufmerksam gemacht. Mehrstellige Relationen sind mit den oben beschriebenen modellbasierten Ansätzen realisierbar. So ist es nicht möglich, den Satz "John gibt Mary das Buch" lediglich durch Vernetzung der Knoten "John", "Mary" und "Buch" darzustellen. Erst durch Einführung eines Gebens-Ereignis-Knotens und Rollenzuweisungen an die Partizipanten des Ereignisses wird ein Netz draus (siehe auch Kranich-Beispiel auf Seite 301).

Vielleicht viel wichtiger als die Frage nach der Brauchbarkeit oder Adäquatheit bestimmter Eigenschaften von Netzwerken ist die Frage, inwieweit hiermit schädliche Artefakte in den Prozeß der Textinterpretation eingeführt werden. So erhebt Busse (1992, S. 111) in einer kritischen Betrachtung der psycholinguistischen Auffassung einer propositionalen Textbasis den Vorwurf eines semantischen Atomismus: Bedeutungselemente, die nicht in Form von Konzepten und Relationen beschrieben werden, blieben unanalysiert. Die sprachlich vorliegende Textform sei so zum bloßen Oberflächenphänomen ohne weitere Bedeutung degradiert.

Obwohl sich diese Kritik sich auf eine Theorie der inneren Repräsentation von Wissen bezieht und damit nur mittelbar relevant für den von uns beschriebenen Einsatz strukturierter Methoden als Arbeitstechnik ist, sollte dieser Aspekt bei der Benutzung von Netzwerken, wie bei jedem Werkzeuggebrauch schlechthin, berücksichtigt werden.

Literatur

- Anderson, J.R. (1989 (2. Auflage)): Kognitive Psychologie: Eine Einführung. Heidelberg: Spektrum der Wissenschaft.
- Anjewierden, A. (1987): Knowledge Acquisition Tools. AICOM, 0(1), 29-38.
- Barrett, E. (Ed). (1989): The Society of Text - Hypertext, Hypermedia and the Social Construction of Information. Cambridge: MIT Press.
- Bogaschewsky, R. (1992): Hypertext-/Hypermedia-Systeme - Ein Überblick. Informatik Spektrum. Berlin: Springer Verlag, 15 (3), 127-143.
- Brachman, R.J., Schmolze, J.G. (1985): An Overview of the KL-ONE Knowledge Representation System. In Cognitive Science (Vol. 9, pp. 207-221).
- Brassel, K.-H. (1992): Zettel - ein flexibles Informationssystem zur Unterstützung kreativer Denktätigkeiten (Diplomarbeit) Universität Koblenz-Lindau, FB Informatik.

- Breutmann, B., Burkhardt, R. (1992):** Objektorientierte Systeme - Grundlagen, Werkzeuge, Einsatz. München: Carl Hanser Verlag.
- Broad, W.J. (1993):** Doing Science on the Network: A Long Way from Gutenberg. New York: New York Times, Science Times, 18.5.1993.
- Budde, R., Kühlenkamp, K., Sylla, K.H., Züllighoven, H. (1989):** Der Entwurf objektorientierter Systeme. HMD, 145, 13-23.
- Burkart, M. (1988):** Neue Thesaurusansätze - frischer Wind in alten Segeln. Nachr. Dok., 39, 207-208.
- Busse, D. (1992):** Textinterpretation. Opladen: Westdeutscher Verlag GmbH.
- Buzan, T. (1984):** Kopftaining. München: Goldmann.
- Chaffin, R., Herrmann, D.J. (1988):** The nature of semantic relations: a comparison of two approaches. In Evens, M.W. (Ed.), Relational Models of the Lexicon (pp. 335-372). Cambridge: Cambridge University Press.
- Dey, I. (1992):** Developing Links - Reducing Fragmentation in Qualitative Research (Vortrag auf der Conference on The Qualitative Research Process and Computing). Bremen:.
- Dijk, T.A.v., Kintsch, W. (1983):** Strategies of Discourse Comprehension. New York: Academic Press.
- Dörner, D. (1989):** Die Logik des Mißlingens: Strategisches Denken in komplexen Situationen. Hamburg: Rowohlt.
- Dreyfus, H.L. (1987):** Künstliche Intelligenz. Von den Grenzen der Denkmaschine und dem Wert der Intuition. Hamburg: Rowohlt.
- Dreyfus, H.L. (1990):** Hermeneutik und Künstliche Intelligenz (Vortrag im Rahmen des ATLAS-Projekts, TU Berlin).
- Eco, U. (1990):** Lector in Fabula - Die Mitarbeit der Interpretation in erzählenden Texten. München: DTV.
- Eigen, M., Winkler, R. (1975):** Das Spiel. Naturgesetze steuern den Zufall. München: Piper Verlag.
- Flusser, V. (1992):** Die Schrift. Frankfurt am Main: Fischer Taschenbuch Verlag GmbH.
- Fuhr, N. (1992):** Konzepte zur Gestaltung zukünftiger Information-Retrieval-Systeme. In: Kühlen, R. (Hrsg.): Experimentelles und praktisches Information Retrieval. Konstanz: Universitätsverlag Konstanz.
- Frawley, W. (1988):** Relational Models and Metascience. In Evens, M.W. (Ed.), Relational Models of the Lexicon (pp. 335-372). Cambridge: Cambridge University Press.
- Glaser, B., Strauss, A.L. (1967):** The Discovery of Grounded Theory. Chicago: Aldine Publishing Co.
- Glaser, B. (1992):** Emergency vs. Forcing: Basics of Grounded Theory Analysis. Mill Valley: Sociology Press.
- Kühlen, R. (1991):** Hypertext - Ein nicht-lineares Medium zwischen Buch und Wissensbank. Berlin: Springer-Verlag.
- Legewie, H. (1989):** Basisoperationen des Kodierens. ATLAS - Internes Arbeitspapier.
- Lindsay, P.H., Norman, D.A. (1981):** Einführung in die Psychologie. Berlin: Springer Verlag.
- Muhr, Th. (1989):** A Computer Aided Textbank- and Textinterpretation System for Qualitative Research in Technology Assessment (ATLAS). 8th International Human Science Research Conference. Aarhus:.

- Muhr, Th. (1991):** ATLAS/ti - A Prototype for the Support of Text Interpretation. In Tesch, R. (Ed.), *Qualitative Sociology* (Vol. 14, pp. 349-371). New York: Human Science Press
- Muhr, Th. (1992):** Catching Bugs & Butterflies in Networks - Using Structured Representations for Qualitative Research (Vortrag auf der Conference on The Qualitative Research Process and Computing, October 7 - 9). Bremen:.
- Nilsson, N.J. (1982):** Artificial Intelligence. Berlin: Springer Verlag.
- Norris, F.R. (1985):** Discrete Structures - an Introduction to Mathematics for Computer Science. New Jersey: Prentice-Hall.
- Parsaye, K., Chignell, M., Khoshafian, S., Wong, H. (1989):** Intelligent Databases - Object-Oriented, Deductive Hypermedia Technologies. New York: Wiley.
- Quillian, M.R. (1968):** Semantic Memory. In Marvin Minsky (Ed.), *Semantic Information Processing* (pp. 217-270). Cambridge: MIT Press.
- Rada, R. (1991):** Hypertext: From Text to Expertext. London: McGraw-Hill.
- Richards, T., Richards, L. (1994):** Using Computers in Qualitative Research. In Denzin, N.K., Lincoln, Y.S. (Ed.), *Handbook of Qualitative Research*. London: SAGE Publications.
- Seibt, V. (1992):** Grafische Darstellung und Manipulation komplexer Datenstrukturen - Alternative Layouts für Graphen. (Studienarbeit am FB Informatik). Berlin: Technische Universität Berlin.
- Seibt, V. (1993):** Modellbasierte Ansätze für die computerunterstützte Textinterpretation (Diplomarbeit am FB Informatik). Berlin: Technische Universität Berlin.
- Sowa, J.F. (1988):** Using a Lexicon of Canonical Graphs in a Semantic Interpreter. In Evens, M.W. (Ed.), *Relational Models of the Lexicon* (pp. 113-137). Cambridge: Cambridge University Press.
- Strauss, A., Corbin, J. (1990):** Basics of Qualitative Research: Grounded Theory Procedures and Techniques. Newbury Park: SAGE Publications.
- Trabant, J. (1989):** Zeichen des Menschen - Elemente der Semiotik. Frankfurt/Main: Fischer.
- Volpert, W. (1980):** Psychologische Handlungstheorie - Anmerkungen zu Stand und Perspektive. In W. Volpert (Ed.), *Beiträge zur Psychologischen Handlungstheorie* (pp. 13-27). Bern: Verlag Hans Huber.
- Werder, Lutz v. (1992):** Kreatives Schreiben in den Wissenschaften. Milow: Schibri-Verlag.
- Wielinga, B.J., Breuker, J.A. (1986):** Models of Expertise. *Proceedings of the ECAI*. Brighton:., 306-318.
- Wielinga, B., Van de Velde, W., Schreiber, G. (1992):** The CommonKADS Framework for Knowledge Modelling (Research Report KADS-II project) ESPRIT Programme of the Commission of the European Communities, Project 5248.

Woods, W.A. (1975): What's in a link: Foundations for Semantic Networks. In Bobrow & Collins (Eds.), *Representation & Understanding* (pp. 35-82). New York: Academic Press.